

# The frequency and rate of pilin antigenic variation in *Neisseria gonorrhoeae*

Alison K. Criss, Kimberly A. Kline<sup>†</sup> and  
H. Steven Seifert\*

Department of Microbiology-Immunology, Northwestern  
University Feinberg School of Medicine, Chicago, IL  
60611, USA.

## Summary

**The pilin antigenic variation (Av) system of *Neisseria gonorrhoeae* (Gc) mediates unidirectional DNA recombination from silent gene copies into the pilin expression locus. A DNA sequencing assay was developed to accurately measure pilin Av in a population of Gc strain FA1090 arising from a defined pilin progenitor under non-selective culture conditions. This assay employs a piliated parental Gc variant with a *recA* allele whose promoter is replaced by *lac*-regulatory elements, allowing for controlled induction of pilin Av. From this assay, the frequency of pilin Av was measured as 0.13 recombination events per cell, with a corresponding rate of pilin Av of  $4 \times 10^{-3}$  events per cell per generation. Most pilin variants retained the parental piliation phenotype, providing the first comprehensive analysis of piliated variants arising from a piliated progenitor. Sequence analysis of pilin variants revealed that a subset of possible recombination events predominated, which differed between piliated and non-piliated progeny. Pilin Av exhibits the highest reported frequency of any pathogenic gene conversion system and can account for the extensive pilin variation detected during human infection.**

## Introduction

Homologous recombination of DNA is a fundamental cellular process involved in replication, DNA repair and the generation of genetic diversity. Gene conversion, the non-reciprocal incorporation of DNA sequence into a homologous locus, underlies immunoglobulin variation in some mammals (Knight, 1992) and is also used by

pathogens such as *Borrelia*, *Campylobacter*, *Plasmodium* and *Trypanosoma* to antigenically vary virulence-associated surface structures (van der Woude and Baumber, 2004). Antigenic variation (Av) has been implicated in evasion of the host humoral immune response and adaptation to new intracellular niches or to different host species (Smith *et al.*, 1995; Liang *et al.*, 2004). During infection, frequencies of gene conversion range from  $10^{-1}$  to  $< 10^{-4}$  recombinants per cell, with corresponding rates of  $10^{-2}$  to  $< 10^{-5}$  recombinants per cell per generation (Roberts *et al.*, 1992; Turner, 1997; Roske *et al.*, 2001; Keely *et al.*, 2003; Tu *et al.*, 2003; Horrocks *et al.*, 2004), but some of these frequencies may reflect selection in the host.

Pilin Av in the human-specific pathogen *Neisseria gonorrhoeae* (Gc) is one of the best-understood systems of gene conversion (Kline *et al.*, 2003). Pilin is the major subunit of the neisserial type IV pili, which are present on all clinical isolates of Gc and are essential for establishing urogenital infection (Cohen and Cannon, 1999). Gc carry a pilin expression locus, *pilE*, and multiple storage loci containing unexpressed pilin coding information, collectively termed *pilS*, on their chromosome. During pilin Av, part or all of a *pilS* copy is transferred unidirectionally into *pilE*, giving rise to new pilin variants that differ in their ability to be expressed or assembled into pili. As a result, pilin Av confers a spectrum of pilus-dependent colony morphologies on Gc, from heavily piliated (P+) to non-piliated (P-).

In the Gc strain FA1090, 19 unique *pilS* copies reside in six discrete chromosomal loci (Hamrick *et al.*, 2001). All *pilS* copies lack the promoter, ribosomal binding site, and 5' 150 bp found at *pilE* (Haas and Meyer, 1986; Segal *et al.*, 1986; Haas *et al.*, 1992). Based on sequence divergence in the *pilS* repertoire, *pilS* copies are comprised of a 5' semivariable region, corresponding to the central ~250 bp of *pilE*, and a 3' hypervariable loop and tail (HV<sub>L</sub> and HV<sub>T</sub>) of ~100 bp in length. Within these variable regions are short (1–34 bp) regions of sequence conserved among all *pilS* copies and *pilE* thought to provide the homology that drives recombination. As with all homologous recombination pathways, RecA function is essential for pilin Av (Kooimey *et al.*, 1987). Pilin Av also requires a RecF-like pathway (RecQ, RecO, RecR and RecJ), the RecA modulator RecX, the growth regulator RdgC, the Rep helicase and the Holliday junction

Accepted 29 July, 2005. \*For correspondence. E-mail h-seifert@northwestern.edu; Tel. (+1) 312 503 9788; Fax (+1) 312 503 1339. <sup>†</sup>Present address: Department of Molecular Microbiology, Washington University School of Medicine, St Louis, MO 63110, USA.

processing enzymes RecG and RuvB (Mehr and Seifert, 1997; 1998; Mehr *et al.*, 2000; Stohl and Seifert, 2001; Skaar *et al.*, 2002; Kline and Seifert, 2005a; Sechman *et al.*, 2005).

Estimates of the frequency of pilin Av are based on P<sup>-</sup> colony appearance or sequence changes in the *pilE* HV<sub>L</sub> and range from 10<sup>-2</sup> to 10<sup>-4</sup> variants per colony-forming unit (cfu) (Segal *et al.*, 1985; Bergstrom *et al.*, 1986; Koomey *et al.*, 1987; Zhang *et al.*, 1992; Wainwright *et al.*, 1994; Serkin and Seifert, 1998; Rohrer *et al.*, 2005). To directly measure pilin Av, we developed a DNA sequencing assay that non-selectively identifies pilin variants arising from a defined *pilE* progenitor. Pilin Av is controlled by the *lac*-inducible regulation of *recA* expression, which allows for the frequency and rate of this process to be accurately measured. We have measured a frequency of pilin Av for FA1090 variant 1-81-S2 of 0.13 recombination events per cfu and a corresponding rate of  $4.0 \times 10^{-3}$  events per cfu per generation, values that are 5–500 times higher than previously reported for this Gc variant. The large number of P<sup>+</sup> and P<sup>-</sup> pilin antigenic variants analysed allowed for definition of the recombination events underlying pilin Av at a previously unavailable level of detail.

## Results

### *Measuring the frequency and rate of pilin Av with the pilE sequencing assay*

We developed a novel DNA sequencing-based assay to measure the total amount of pilin Av independent of colony morphology changes in a population of Gc. This assay depends on three key factors. First, Gc of strain background FA1090 were used, as all *pilS* copies in this strain have been sequenced (Hamrick *et al.*, 2001). Second, the specific FA1090 isolate used harbours the *recA6* allele, in which RecA expression is controlled by *lac*-regulatory elements (Seifert, 1997). In the *recA6* genetic background, RecA-dependent processes like pilin Av are initiated only when Gc are grown in the presence of isopropyl β-D-thiogalactopyranoside (IPTG) (Seifert, 1997; Serkin and Seifert, 1998; Rohrer *et al.*, 2005). The *recA6* isolate encodes the 1-81-S2 pilin variant (Seifert *et al.*, 1994), a highly pilated variant that has been used previously to measure pilin Av (Serkin and Seifert, 1998; Rohrer *et al.*, 2005). The combination of the *recA6* allele with the 1-81-S2 variant sequence ensures that the parental colonies uniformly express one pilin protein. Third, large-scale DNA sequencing allowed the *pilE* genes from a large number of randomly selected progeny to be rapidly examined. In combination, these factors yield the most complete analysis of pilin Av to date.

To perform the assay, three to seven single colonies of FA1090 1-81-S2 *recA6* Gc are grown in the presence of IPTG to allow pilin Av to occur. These colonies are dispersed and grown on solid medium in the absence of IPTG. At least 48 cfu arising from each parental colony are isolated and serially passaged to ensure the recovery of pure populations of Gc. The *pilE* loci of the progeny are sequenced and compared with both the parental *pilE* allele and the *pilS* copies present in the FA1090 genome. These numbers of starting and progeny colonies were determined statistically to yield accurate, reproducible measurements of pilin Av (see *Experimental procedures*). The assay was performed with three independent cultures of Gc, generating a total of 497 progeny colonies that were examined for *pilE* sequence changes.

From 5.2 to 17.3% of *pilE* clones in the three sequenced populations differed from 1-81-S2, averaging  $11.8 \pm 3.5\%$  variation. Strikingly, 0.96% of the *pilE* sequences in each replicate population had undergone two recombination events, comprising 13.3% of all non-parental *pilE* sequences. A recombination event is defined as the incorporation of new *pilS* sequence into *pilE*; when multiple recombination events have occurred, sequences corresponding to different *pilS* copies appear at various sites in *pilE*. Detection of two recombination events in a *pilE* sequence was a reasonable expectation: if 10% of clones exhibit one event, 1% would on average exhibit two. However, multiple recombination events per *pilE* sequence have never before been taken into account when measuring pilin Av. As the sequencing assay had the unique ability to identify these events, we redefined the frequency of pilin Av as the number of recombination events measured in a population divided by the total number of *pilE* genes sequenced, i.e. the number of cfu. In the 497 P<sup>+</sup> cfu examined, the frequency of pilin Av was  $0.13 \pm 0.04$  recombination events per cfu. The DNA sequencing assay is therefore able to detect substantial amounts of pilin Av in a Gc population cultured under non-selective conditions.

Assuming each colony grown in the presence of IPTG arose from one cfu and contained 19 generations of bacteria ( $\sim 3 \times 10^5$  cfu per colony), the resulting rate of pilin Av is  $\sim 6.8 \times 10^{-3}$  recombination events per cfu per generation. In order to more accurately measure the rate of pilin Av, FA1090 1-81-S2 *recA6* Gc were grown in liquid medium containing IPTG, and the frequency of pilin Av was analysed at 3 h intervals up to 12 h after IPTG induction for three independent cultures. The average generation time of these cultures was 61 min. No antigenic variants were detected in any of the replicate cultures at 3 h after IPTG induction (a frequency of  $< 9.5 \times 10^{-3}$  events per cfu). At 6 h, the average frequency was  $1.1 \times 10^{-2}$  events per cfu; at 9 h,  $3.5 \times 10^{-2}$  events per cfu; at 12 h,  $5.1 \times 10^{-2}$  events per cfu. Using the data from the

6–12 h time points, the rate of pilin Av was calculated to be  $4.0 \pm 0.6 \times 10^{-3}$  recombination events per cfu per generation.

The results with FA1090 1-81-S2 *recA6* were validated using Gc with a wild-type *recA* allele, in which pilin Av occurs constitutively. Three experiments were conducted, yielding 180 *pilE* clones. Of progeny cfu,  $22.5 \pm 2.7\%$  encoded non-parental *pilE* sequences, with a corresponding frequency of pilin Av of  $0.25 \pm 0.01$  recombination events per cfu. This frequency is within twofold of that measured in *recA6* Gc. The rate of pilin Av of Gc with a wild-type *recA* allele was estimated by culturing Gc in liquid medium and performing the sequencing assay at 0, 3, 6, 9 and 12 h after inoculation. As in this experiment the starting inoculum exhibited a frequency of pilin Av of 0.25 events per cfu, the rate was calculated as the change in frequency at each time point divided by the number of generations and was found to be  $1.8 \pm 0.4 \times 10^{-2}$  recombination events per cfu per generation. The frequency and rate calculations with this Gc strain are substantially higher than measured in *recA6* Gc, but these values should only be considered estimates as the starting colonies of FA1090 1-81-S2 Gc may have contained bacteria with non-parental *pilE* sequence. Furthermore, it is likely that each wild-type *recA* colony contains a different percentage of non-parental *pilE* variants, but as these variants arise stochastically, the basal level of variation in a colony cannot be accurately assessed. When the frequency of pilin Av is measured, the variant sequences present at the start are recorded as if they had arisen during the course of the experiment (although they were already present at the start), thus affecting the calculated rate. In contrast, at the time of *recA* induction the starting population of the *recA6* Gc is uniform in its *pilE* sequence, which then produces a defined rate. These results show that the values calculated with Gc with a wild-type *recA* allele are similar to those measured in the *recA6* strain, and even after taking into account the variability inherent in wild-type colonies, may in fact be higher. We conclude that the *pilE* sequencing assay for Gc reproducibly measures pilin Av, and it reveals that the frequency and rate of this process are significantly higher than previously reported (Serkin and Seifert, 1998; Rohrer *et al.*, 2005).

#### Non-random incorporation of *pilS* sequences in P+ *pilE* variants

Before this study, large-scale analysis of pilin antigenic variants was limited to Gc exhibiting pilus-dependent colony morphology changes. However, results from the *pilE* sequencing assays clearly demonstrate that a sizable percentage of P+ Gc arising from a P+ progenitor undergo pilin Av. We therefore used the sequence information from these assays to investigate the recombination events arising

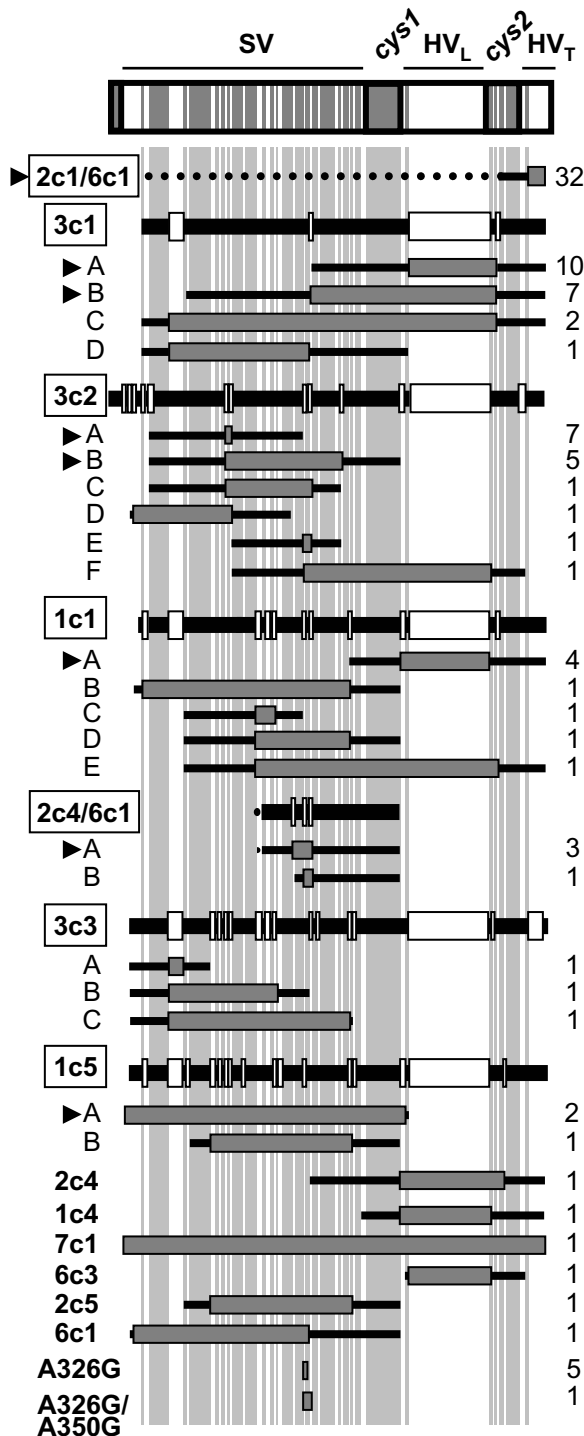
in P+ variants of FA1090 1-81-S2. Ninety-three non-parental *pilE* sequences were recovered from 1-81-S2 harbouring either the *recA6* or the wild-type *recA* allele. Five of the 93 *pilE* clones had undergone two recombination events at *pilE*, generating 98 total recombination events. Ninety-two of the 98 events could be unambiguously assigned to one or two *pilS* copies. The remainder diverged from 1-81-S2 at 1–2 bp and corresponded to sequences found in multiple silent copies. It is likely that these single-nucleotide changes are true antigenic variants and not point mutants, as (i) the nucleotide changes are in variable regions and are present in several *pilS* copies that have flanking sequence identity to 1-81-S2 and (ii) no single-nucleotide changes were detected elsewhere in *pilE*, including conserved sequences.

The 92 P+ recombination events with identifiable donor silent copies had sequence corresponding to 12 of the 19 *pilS* copies in Gc strain FA1090. Within this subset, certain silent copies predominated (Table 1). The most frequently recovered recombinant exhibited a change in the HV<sub>T</sub> region of *pilE* corresponding to the sequence found in both *pilS2c1* and *pilS6c1* (Fig. 1). It cannot be determined whether one or both of these donors contributed to this recombination event, as both share sequence identity with 1-81-S2 *pilE* upstream of the HV<sub>T</sub>. However, *pilS2c1* is identical to 1-81-S2 *pilE* for 268 bp upstream of the HV<sub>T</sub> (dotted line in *2c1/6c1*, Fig. 1), whereas *pilS6c1* only shares 32 bp of sequence with 1-81-S2 *pilE* in this region (solid line). Sequences from *pilS3c1*, *pilS1c1* and *pilS3c2*

**Table 1.** Distribution of *pilS* copies in variants of FA1090 1-81-S2 Gc.

<i>pilS</i> copy	% appearance in:	
	P+	P–
<i>pilS2c1</i> OR <i>pilS6c1</i>	34.8	6.6
<i>pilS3c1</i>	21.7	1.6
<i>pilS3c2</i>	17.4	36.1
<i>pilS1c1</i>	8.3	–
<i>pilS2c4</i> OR <i>pilS6c1</i>	4.3	–
<i>pilS3c3</i>	3.3	–
<i>pilS1c5</i>	3.3	1.6
<i>pilS2c4</i>	1.1	36.1
<i>pilS1c4</i>	1.1	4.9
<i>pilS7c1</i>	1.1	1.6
<i>pilS6c3</i>	1.1	–
<i>pilS2c5</i>	1.1	–
<i>pilS6c1</i>	1.1	3.3
<i>pilS2c1</i>	–	4.9
<i>pilS6c2</i>	–	1.6
<i>pilS1c3</i>	–	1.6

P+ and P– variants were isolated after passage of FA1090 1-81-S2 and its *recA6* derivative, yielding 92 P+ and 61 P– recombination events where the donor silent copy could be assigned to one or two *pilS* sequences. The per cent appearance of each *pilS* copy, or sequence corresponding to two *pilS* copies, was determined. Each *pilS* copy is annotated for its copy number and position in *pilS* as follows: *pilS3c1* = *pilS* locus 3, copy 1.



also appeared repeatedly. Notably, four recombination events, *pilS2c1/6c1*, *pilS3c1-A* and *-B*, and *pilS3c2-A*, accounted for 57% of the events detected (Fig. 1), and all arose in more than one of the six *pilE* sequencing assays. In total, eight *pilE* variants emerged in more than one of the sequencing assays performed (arrowheads, Fig. 1),

**Fig. 1.** Recombination events in P+ variants of Gc strain FA1090. *pilS* copies are abbreviated as in Table 1. The positions in *pilE* of the semivariable (SV), hypervariable loop (HV<sub>L</sub>) and tail (HV<sub>T</sub>), and conserved *cys1* and *cys2* sequences are indicated. Sequences conserved among all *pilS* copies are grey and divergent regions are white; both are to scale. Recombination events involving the same donor *pilS* copy are grouped together and are presented in decreasing order of frequency of appearance in *pilE*. Grey boxes indicate recombinant *pilS* sequence in 1-81-S2, and flanking sequence shared by 1-81-S2 and the donor *pilS* copy is indicated by flanking black lines. Six *pilS* copies (boxed) participated in more than one recombination event at *pilE* (labelled A, B, etc.). The alignment of each *pilS* copy with the 1-81-S2 sequence is shown above the corresponding recombination events; regions of each *pilS* copy that share identity with 1-81-S2 are indicated by black bars and divergent sequences by white boxes. In recombination events where two *pilS* copies could have served as donor (*2c4/6c1* and *2c1/6c1*), flanking sequence shared with both copies is indicated with a black line; sequence shared with only one copy is indicated by a dotted line. In *pilS2c1/6c1*, the dotted line represents additional homology between *pilS2c1* and 1-81-S2 not present in *pilS6c1*; in *pilS2c4*, the dotted line represents additional homology between *pilS6c1* and 1-81-S2 not present in *pilS2c4*. The number of times each recombination event was recovered is indicated on the right. The eight pilin variants recovered in independent cultures of Gc are indicated by arrowheads. A326G and A326G/A350G are sequences found in multiple *pilS* copies.

indicating that certain recombination events are more frequently recovered than others. The most predominant *pilS* copies appeared at similar levels in both *recA6* and wild-type *recA* genetic backgrounds, indicating there is no mechanistic difference in pilin Av between these two strains (data not shown). The relative frequency of incorporation of *pilS* copies into the *pilE* HV<sub>L</sub> differed from full-length *pilE*, but correlated with a previous examination of the repertoire of donor *pilS* sequences in the HV<sub>L</sub> arising from the 1-81-S2 variant (Rohrer *et al.*, 2005).

All 98 P+ recombination events were examined for the length and position of variant sequence in *pilE* (Fig. 1). The length of contiguous variant sequence incorporated in *pilE* averaged 55 bp and ranged from 1 bp (A326G) to greater than 200 bp (*pilS3c1* recombinant C, *pilS1c5-B*). Notably, 50% of the recombination events involved incorporation of 15 bp or less of *pilS* sequence (*pilS2c1/6c1*, *pilS3c2-A* and *-E*, *pilS3c3-A*). In contrast, only one recombination event involved incorporation of an entire *pilS* copy (*pilS7c1*). The 98 P+ recombination events occurred throughout *pilE*, with 50% of variants exhibiting changes in the semivariable region. As previously observed (Seifert, 1996; Howell-Adams and Seifert, 2000; Hamrick *et al.*, 2001), the location of the end points of variant information did not support the minicassette model of pilin Av (Haas and Meyer, 1986). Taken together, these results indicate that P+ antigenic variants derived from 1-81-S2 *pilE* are generated from a subset of *pilS* copies that donate short stretches of sequence to *pilE*, and that a small number of pilin variants exhibiting identical recombination events at *pilE* arise repeatedly from this progenitor pilin sequence.

Generation of P<sup>-</sup> variants by pilin Av

*Neisseria gonorrhoeae* colonies of P<sup>-</sup> morphology arise from pilin Av, L-pilin production (the incorporation of multiple *pilS* copies in tandem into *pilE*, yielding an aberrantly long, non-functional protein), deletion of the *pilE* locus and phase variation of the structural PilC proteins (Hagblom *et al.*, 1985; Segal *et al.*, 1985; Swanson *et al.*, 1985; 1986; Jonsson *et al.*, 1991). To explore the relative contribution of these events to P<sup>-</sup> colony formation, FA1090 1-81-S2 *recA6* was grown on solid medium containing IPTG to induce pilin Av, sectors of colonies exhibiting a less piliated morphology were isolated, and the *pilE* genes were sequenced from progeny retaining the P<sup>-</sup> morphology. Three replicate experiments were performed with a total of 67 P<sup>-</sup> cfu analysed.

The appearance of P<sup>-</sup> colonies in a population has been used as a surrogate measure of pilin Av, and we observed that 58.6% of the P<sup>-</sup> clones had undergone pilin Av. The remaining 41.4% of P<sup>-</sup> colonies exhibited one of three phenotypes: (i) production of L-pilin, determined by the generation of an abnormally long polymerase chain reaction (PCR) product using *pilE*-specific primers (6.5% of the P<sup>-</sup> isolates), (ii) deletion of the *pilE* locus, assessed by the inability to produce a PCR product with these same *pilE*-specific primers (20.2%) and (iii) PilC phase variation, or the retention of parental *pilE* sequence while exhibiting P<sup>-</sup> morphology (14.7%). The absence of PilC expression was confirmed by Western blot (data not shown). Two of the 67 P<sup>-</sup> cfu exhibited sequence changes that were detected in the P<sup>+</sup> population and were also PilC phase variants (*2c1/6c1* and *1c5*, Fig. 2). These results demonstrate that pilin Av is a common mechanism for generating P<sup>-</sup> variants, but other mechanisms are of equal prevalence and should be taken into consideration when using phase variation as a surrogate measure of pilin Av.

The 47 *recA6* and five wild-type *recA* P<sup>-</sup> pilin antigenic variants yielded a total of 52 P<sup>-</sup> variants and 64 recombination events. Sixty-one of the 64 recombination events could be assigned to one or two *pilS* copies (Fig. 2). Compared with the P<sup>+</sup> variants, a smaller and different subset of *pilS* copies were represented in the P<sup>-</sup> variants (Table 1). As with the P<sup>+</sup> variants, the most predominant *pilS* copies appeared in each of the replicate P<sup>-</sup> populations at a similarly high frequency (data not shown). These variants fell into two categories. The first class of variants, accounting for 7% of the P<sup>-</sup> sequences, was predicted to produce full-length pilin protein but produced a P<sup>-</sup> colony phenotype. These variants are classified as 'assembly mutants', where pilin monomers are not effectively assembled into the growing pilus (Swanson *et al.*, 1986). No pilin protein was detected in these variants by Western blot,

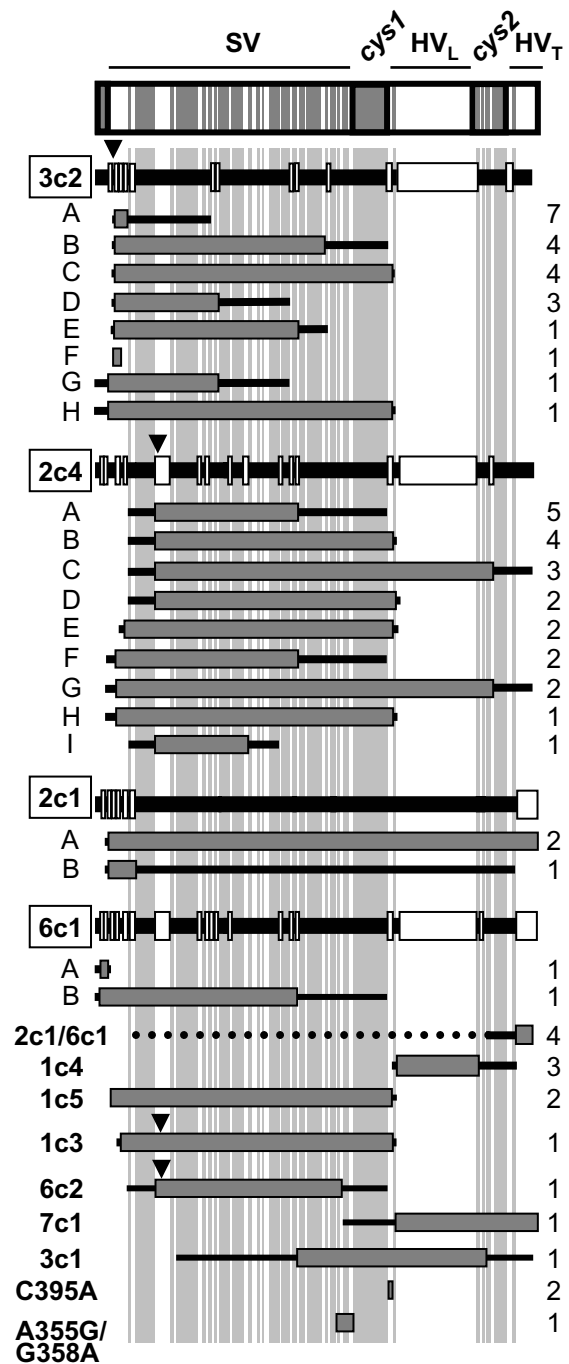


Fig. 2. Recombination events in P<sup>-</sup> variants of Gc strain FA1090. Labelling of divergent and conserved regions, and alignments of *pilS* copies with *pilE*, is as in Fig. 1. C395A and A355G/G358A are sequences found in multiple *pilS* copies. Black arrowheads indicate the position of frameshifts that are predicted to cause premature truncation of pilin.

but PilC was still expressed (data not shown). This class of variants incorporated *pilS2c1* or *pilS6c1* in the semi-variable region of *pilE* (*2c1*-A and -B and *6c1*-A and -B, Fig. 2).

```

1-81-S2 CGTCACCGAGTATTACCTGAATCACGGCATATGGCCGAA
pils3c2 -C-----A-----A*-----GA-----C--
pils2c4 G--TG---G---G-----G-----GA-----G-
pils6c2 ---TG---G---C-----A-----G-
pils1c3 ---TG---G---G-C-----A-----G-

1-81-S2 AGACAACACTTCTGCGGGCGTGGCATCC*GCTTCA*ACA
pils3c2 -----GGC-----*-----*-----
pils2c4 -----*C-CC-*C-G
pils6c2 -----*C-CC-T*C-G
pils1c3 -A-----C-----CC-TCC-G

1-81-S2 **ATCAAAGGCCAAATATGTTTCAGAAAGTTGAAGTCGCAA
pils3c2 **-----A---
pils2c4 (AC)-----A-A-GC---ACG-----
pils6c2 (AC)-----A-GC---ACG-----
pils1c3 AC-----A---(A)-----

1-81-S2 AAGGCGTCGTTACCGCCAAATGGCTTCAACCGCGTAA
pils3c2 -C-----G-----G-----
pils2c4 -----C-----G-----
pils6c2 -C-----G-----AAA---GA-----

```

**Fig. 3.** *pilS* copies in strain FA1090 that produce truncated pilin proteins and confer P<sup>-</sup> colony morphology. Alignments of *pilS* copies with 1-81-S2 *pilE* begin at nucleotide 159 of *pilE*, numbered from the first nucleotide of the ATG start codon. Sequence divergences from 1-81-S2 *pilE* are indicated; identical nucleotide to 1-81-S2 (-); nucleotide missing relative to other *pilS* copies or *pilE* (\*). Frameshifts relative to 1-81-S2 are circled. Each *pilS* sequence terminates with the predicted stop codon (boxed).

The majority of P<sup>-</sup> pilin variants also failed to react with pilin-specific antiserum, but these variants were predicted to encode truncated proteins. These P<sup>-</sup> variants contained sequence encoded in one of four *pilS* copies: *pilS3c2*, *pilS2c4*, *pilS6c2* and *pilS1c3* (Fig. 3). Incorporation of *pilS2c4*, *pilS3c2* and *pilS6c2* at the 5' end of the *pilE* semivariable region generates a stop codon after residue 94 or 95, and incorporation of *pilS1c3* introduces a stop codon at position 78. Three of the four *pilS* copies contain poly(C) tracts that could vary in length as a result of replication errors to restore a P<sup>+</sup> phenotype (Koohey *et al.*, 1987). As with P<sup>+</sup> variants, certain *pilS* sequences in these P<sup>-</sup> variants predominated, with *pilS2c4* and *pilS3c2* detected more frequently than the other two (Table 1).

Compared with P<sup>+</sup> variants, P<sup>-</sup> variants on average contained longer tracts of *pilS* sequence which was incorporated closer to the 5' end of the *pilE* semivariable region (Fig. 2). Eighty-six per cent of P<sup>-</sup> recombination events occurred in the semivariable region of *pilE*, and the average minimum length of *pilS* sequence incorporation in P<sup>-</sup> variants was 133 bp. Although *pilS* incorporation as short as 1 bp could be detected (C395A, Fig. 2), only 15.7% of variants had incorporated 15 bp or less of *pilS* sequence. In summary, pilin Av accounts for the majority of P<sup>-</sup> colonies generated after *in vitro* passage of Gc, and these variants arise from recombination with a subset of *pilS* copies that encode truncated or poorly assembled pilin proteins.

### Minimal sequence identity is required for pilin Av

RecA-dependent homologous recombination requires sequence identity between donor and recipient DNA of at least 73 bp (Nash, 1996). Each donor *pilS* copy identified in the variant *pilE* population was aligned with 1-81-S2 *pilE* to identify regions of shared sequence (depicted for a subset of *pilS* copies in Figs 1 and 2). For each recombination event where the donor *pilS* copy could be unambiguously identified, the length of identical sequence between the incoming copy and 1-81-S2 *pilE* flanking the recombinant tract was measured. The median amount of shared sequence was 44 bp at the 5' flanking region and 39 bp at the 3' flanking region, and eight independent recombination events were detected where shared sequence identity was 2–4 bp. No significant difference was noted in the amount of sequence overlap between P<sup>+</sup> and P<sup>-</sup> variants. Minimal regions of shared sequence identity were not necessarily paired with especially extensive sequence identity at the other end of the recombination tract, and the length of sequence overlap was not related to the length of *pilS* sequence appearing in *pilE*. This analysis confirms that minimal sequence identity is needed for *pilE* recombination (Howell-Adams and Seifert, 2000) and further demonstrates that recombination is independent of the length or identity of the *pilS* sequence incorporated, lending further support for the current understanding of pilin Av as a specialized form of homologous recombination.

### Discussion

Gonococcal pilin Av is a prototype for prokaryotic gene conversion systems, yet a direct measure of the frequency and rate of this process has never been reported. We developed a *pilE* sequencing assay to address this issue, which revealed that one out of four to eight cfu in a population has undergone recombination at *pilE*. These frequencies are 3- to 10-fold higher than measured with a reverse transcription polymerase chain reaction (RT-PCR) assay examining variants in the HV<sub>L</sub>, the most complete assay for pilin Av until this point (Serkin and Seifert, 1998; Rohrer *et al.*, 2005). Strikingly, the maximum frequency of pilin Av may be even higher than reported here, as FA1090 Gc with a different starting *pilE* sequence was found to produce a greater percentage of HV<sub>L</sub> variants than 1-81-S2 (Rohrer *et al.*, 2005). To our knowledge, pilin Av exhibits the highest frequency of gene conversion of any pathogenic Av system examined to this point (van der Woude and Baumler, 2004). One drawback to this assay is its cost, considering that a minimum of 480 PCR products need to be sequenced per time point for a given strain. However, this assay demonstrated that most antigenic variants retain the parental P<sup>+</sup> phenotype and pro-

vided the first comprehensive analysis of P<sup>+</sup> variants arising from a P<sup>+</sup> progenitor. It also identified a novel means by which P<sup>-</sup> variants arise from a P<sup>+</sup> progenitor. Finally, by generating 159 non-parental *pilE* sequences encompassing 48 unique pilin variants (Fig. S1 in *Supplementary material*), the assay produced the most complete picture available about the recombination events of pilin Av.

It is apparent from the P<sup>+</sup> and P<sup>-</sup> variant *pilE* sequences arising from FA1090 1-81-S2 that certain *pilS* copies predominate, even in the absence of any apparent selection pressure. We have calculated that 963 independent antigenic variants could arise following one recombination event with 1-81-S2 *pilE*, if at least 2 bp of sequence identity exists at the end of a tract of new *pilE* sequence. This is inconsistent with the observation that only three recombination events account for 50% of the P<sup>+</sup> variant population (*2c1/6c1* and *3c1-A* and *-B*, Fig. 1). These are not isolated events, as these and other high-frequency recombinants occurred in multiple experiments (arrowheads, Fig. 1). Several explanations for the non-random distribution of *pilS* donors were considered, including the chromosomal location of a *pilS* copy or its position in the *pilS* locus, its relative percentage of sequence identity to *pilE*, the presence of pilin-associated repeat elements (Hamrick *et al.*, 2001) and the length of the sequence tracts flanking the *pilE/pilS* recombinants. The Av data presented in this work do not individually support any of these possibilities. We postulate that either a combination of these factors dictates the prevalence of *pilE* recombination events, or some as yet unknown sequence element or bacterial factor influences the frequency of *pilS* recombination at the *pilE* locus.

In FA1090 1-81-S2, the predominant mechanism for generating P<sup>-</sup> variants was recombination with silent copies that shift the *pilE* coding frame and prematurely incorporate a stop codon. These truncations were predicted from the sequence analysis of *pilS* copies by Hamrick *et al.* (2001) but were not previously recovered in P<sup>-</sup> Gc. It is important to note that these variants differ from S-pilin variants, which encode full-length pilin proteins that undergo post-translational removal of their N-termini (Kooimey *et al.*, 1991). Although we have no direct evidence that P<sup>-</sup> variants occur at a measurable frequency during infection, the retention of *pilS* copies encoding truncated pilins in the genome suggests that non-piliated variants may have a role in infection. The first possibility for such a role is that non-piliated Gc have a growth advantage *in vivo* as, *in vitro*, non-piliated Gc have an enhanced metabolic rate compared with an isogenic pilated strain and can outgrow their pilated counterparts when cultured together in liquid medium (H.S. Seifert, unpubl. results). The second possibility is that loss of piliation is advantageous during infection. For

instance, the reduced adherence capacity of non-piliated bacteria may facilitate Gc movement within a person or transfer between individuals. Furthermore, P<sup>-</sup> Gc would not be recognized by host pilin-specific antibodies, and in this light a lack of piliation can be considered an extreme form of Av. The third possibility is that non-piliated Gc may alter host cell signalling. In epithelial cells, neisserial pili and their retractile forces induce Ca<sup>2+</sup> spikes and the activation of protein kinases (Kallstrom *et al.*, 1998; Howie *et al.*, 2005); modulation of these signals by P<sup>-</sup> Gc could affect, for example, the inflammatory response. If any or all of these possibilities are found to be important for Gc pathogenesis, pilin Av would provide one high-frequency mechanism behind P<sup>+</sup> ↔ P<sup>-</sup> variation.

The extent of pilin Av measured by the DNA sequencing assay can be compared with that reported during human infection. The majority of Gc recovered from experimentally infected male volunteers encode non-parental *pilE* sequences after 24 h of infection, and some *pilE* genes contain more than one recombination tract (Swanson *et al.*, 1987; Seifert *et al.*, 1994; Hamrick *et al.*, 2001). These high levels of pilin variation were thought to be due either to stimulation of the frequency of pilin Av or to selection against the parental pilin protein. In contrast, results from the sequencing assay would argue that the frequency of pilin Av *in vitro* is sufficient to account for all pilin variation that occurs during infection. It is not possible to directly compare the rate of pilin Av measured *in vitro* with that occurring *in vivo* because the number of generations of growth in the host cannot be determined: colonization bottlenecks occur early after inoculation that limit the number of bacteria that actually establish infection (Cohen and Cannon, 1999). Assuming all the progeny arising during acute infection were descended from one parental cfu and an *in vitro* rate of pilin Av, after 1 day approximately 50% of isolates would have undergone a recombination event, which is similar to what has been observed in experimentally infected individuals. Our results also correlate with the observation made in human volunteer studies that a small population of antigenic variants, involving a subset of *pilS* copies, predominate, without needing to invoke selection *in vivo* (Seifert *et al.*, 1994; Hamrick *et al.*, 2001). However, it is likely that in natural, long-term infections, immune selection is important for driving the appearance of a subpopulation of pilin variants. As each newly arising pilin variant would give rise to its own unique repertoire of pilin sequences (Rohrer *et al.*, 2005), this ever-expanding pool of variants could facilitate immune evasion as Gc are transmitted within a sexual network and back to previously exposed individuals. We conclude that all observations made about gonococcal pilin Av *in vitro* are sufficient to explain the repertoire of pilin variants arising during infection, underscoring the

importance of this process to the persistence and spread of Gc in the human population.

## Experimental procedures

### Bacterial strains and growth conditions

1-81-S2 is a P+ variant of Gc strain FA1090 arising after experimental infection of a male volunteer (Seifert *et al.*, 1994). FA1090 1-81-S2 *recA6* contains an IPTG-regulatable Gc *recA* gene, which allows for control of *recA* expression. These strains are phenotypically RecA- when IPTG is absent from the growth medium and phenotypically RecA+ in the presence of IPTG (Seifert, 1997). Gc were grown on GC Medium Base (Difco) plus Kellogg supplements (Kellogg *et al.*, 1963) (GCB) at 37°C in 5% CO<sub>2</sub> for 20 h, with the addition of 1 mM IPTG (Diagnostic Chemicals Limited) when necessary. Colonies were observed in the stereomicroscope to ensure retention of the desired piliation phenotype. Liquid growth of Gc was performed in Gc liquid medium containing Kellogg's supplements and 0.042% Na<sub>2</sub>HCO<sub>3</sub> (GCBL) as described (Serkin and Seifert, 1998).

### Pilin variation assays

**Growth on solid medium.** The pilin variation assay was performed as described (Kline and Seifert, 2005b). Three to seven colonies of FA1090 1-81-S2 *recA6* Gc grown on GCB-IPTG were each passaged on GCB. Approximately 48 P+ colonies that arose from each starter colony were passaged two more times to ensure clonal populations, then stored at -80°C. For Gc with the wild-type *recA* allele, one P+ colony of Gc with 1-81-S2 *pilE* sequence was passaged on GCB. Forty-eight to ninety-six cfu of both P+ and P- colony morphology were collected and stored at -80°C without further passage. P- colonies comprised ~5% of this population.

To examine P- variants in FA1090 1-81-S2 *recA6*, P- outgrowths arising from P+ colonies on GCB-IPTG were isolated and streak-purified for several days on GCB to ensure a homogeneous P- population was obtained. In each experiment, up to 32 independent P- clones were examined.

**Growth in liquid medium.** Liquid-grown FA1090 1-81-S2 *recA6* were diluted into GCBL containing 1 mM IPTG at 37°C. At 3, 6, 9 and 12 h after IPTG induction, an aliquot of the culture was serially diluted and plated. The number of cfu present at each time point was determined, and 96 cfu were collected from each time point. The experiment was performed in triplicate.

**Statistical calculations.** To determine the minimum sample size necessary to yield an accurate measurement of pilin Av, we performed a small pilot study with three 1-81-S2 *recA6* starter colonies and 16–24 sequenced progeny from each, which had a standard deviation of 2.8% (data not shown). The formula

$$N = [(z * s.d.) / E]^2$$

was then applied, where *N* = sample size, *z* = value corresponding to the confidence interval, s.d. = standard deviation

of the sample and *E* = acceptable margin of error for the assay. Using *z* = 95% confidence interval and *E* ≤ 1% margin of error, *N* was calculated at 48 progeny per starter colony.

***pilE* sequence analysis.** The *pilE* locus was amplified by polymerase chain reaction (PCR) from Gc lysates as described previously (Kline and Seifert, 2005b). Sequencing reactions were carried out with the CEQ Dye Terminator Cycle Sequencing Quick Start Kit using primers specific for *pilE* and were run on a CEQ 2000XL automated sequencer (Beckman Coulter) according to the manufacturer's instructions. Alternatively, purified PCR products were sequenced commercially (SeqWright, Houston, TX).

**Calculations for pilin Av.** The *pilE* sequence from each clone was aligned with the parental 1-81-S2 *pilE*, as well as with all *pilS* copies present in the FA1090 genome (Hamrick *et al.*, 2001), using the ALIGNX program of the Vector NTI software (Informax). In all cases, sequence changes observed in the isolates corresponded to one or more *pilS* copies. Non-parental *pilE* sequences were translated using the EDITSEQ program in the Lasergene software package (DNASTar).

Each independent incorporation of *pilS* sequence into a *pilE* clone was designated a recombination event. For FA1090 *recA6* Gc, the frequency of pilin Av was calculated by the following formula: frequency = total number of recombination events at *pilE* ÷ total number of isolates sequenced. For FA1090 1-81-S2 Gc, the frequency of pilin Av was determined separately for P+ and P- cfu, and the overall frequency was calculated by multiplying the frequency for each population by the fraction of colonies with the given colony morphology. Data are presented as the average of the frequencies measured from replicate samples ± SEM. The rate of pilin Av is presented as the average change in frequency divided by the average number of bacterial generations over the course of the experiment ± SEM.

**Analysis of PilC and pilin production in P- variants.** Lysates of P- variants were separated by SDS-polyacrylamide gel electrophoresis (10% polyacrylamide for PilC, 15% for pilin) and transferred to polyvinylidene difluoride (Millipore). PilC was detected using a polyclonal anti-PilC antiserum (gift of J. Pfeifer and S. Normark) (Nassif *et al.*, 1994) and pilin detected using T36 polyclonal antiserum (gift of M. So) (Forest *et al.*, 1996), with both followed by goat anti-rabbit IgG coupled to horseradish peroxidase (Chemicon). Blots were developed using the ECLPlus chemiluminescence reagent (Amersham).

## Acknowledgements

This work was supported by Public Health Service Grants R37 AI033493 and R01 AI044239 to H.S.S. and Ruth L. Kirchstein National Research Service Award F32 AI056681 to A.K.C.

## References

- Bergstrom, S., Robbins, K., Koomey, J.M., and Swanson, J. (1986) Piliation control mechanisms in *Neisseria gonorrhoeae*. *Proc Natl Acad Sci USA* **83**: 3890–3894.

- Cohen, M.S., and Cannon, J.G. (1999) Human experimentation with *Neisseria gonorrhoeae*: progress and goals. *J Infect Dis* **179** (Suppl. 2): S375–S379.
- Forest, K.T., Bernstein, S.L., Getzoff, E.D., So, M., Tribbick, G., Geysen, H.M.X., *et al.* (1996) Assembly and antigenicity of the *Neisseria gonorrhoeae* pilus mapped with antibodies. *Infect Immun* **64**: 644–652.
- Haas, R., and Meyer, T.F. (1986) The repertoire of silent pilus genes in *Neisseria gonorrhoeae*: evidence for gene conversion. *Cell* **44**: 107–115.
- Haas, R., Veit, S., and Meyer, T.F. (1992) Silent pilin genes of *Neisseria gonorrhoeae* MS11 and the occurrence of related hypervariant sequences among other gonococcal isolates. *Mol Microbiol* **6**: 197–208.
- Hagblom, P., Segal, E., Billyard, E., and So, M. (1985) Intragenic recombination leads to pilus antigenic variation in *Neisseria gonorrhoeae*. *Nature* **315**: 156–158.
- Hamrick, T.S., Dempsey, J.A., Cohen, M.S., and Cannon, J.G. (2001) Antigenic variation of gonococcal pilin expression *in vivo*: analysis of the strain FA1090 pilin repertoire and identification of the *pilS* gene copies recombining with *pilE* during experimental human infection. *Microbiology* **147**: 839–849.
- Horrocks, P., Pinches, R., Christodoulou, Z., Kyes, S.A., and Newbold, C.I. (2004) Variable *var* transition rates underlie antigenic variation in malaria. *Proc Natl Acad Sci USA* **101**: 11129–11134.
- Howell-Adams, B., and Seifert, H.S. (2000) Molecular models accounting for the gene conversion reactions mediating gonococcal pilin antigenic variation. *Mol Microbiol* **37**: 1146–1159.
- Howie, H.L., Glogauer, M., and So, M. (2005) The *N. gonorrhoeae* type IV pilus stimulates mechanosensitive pathways and cytoprotection through a *pilT*-dependent mechanism. *PLoS Biol* **3**: e100.
- Jonsson, A.B., Nyberg, G., and Normark, S. (1991) Phase variation of gonococcal pili by frameshift mutation in *pilC*, a novel gene for pilus assembly. *EMBO J* **10**: 477–488.
- Kallstrom, H., Islam, M.S., Berggren, P.O., and Jonsson, A.B. (1998) Cell signaling by the type IV pili of pathogenic *Neisseria*. *J Biol Chem* **273**: 21777–21782.
- Keely, S.P., Cushion, M.T., and Stringer, J.R. (2003) Diversity at the locus associated with transcription of a variable surface antigen of *Pneumocystis carinii* as an index of population structure and dynamics in infected rats. *Infect Immun* **71**: 47–60.
- Kellogg, D.S., Jr, Peacock, W.L., Deacon, W.E., Brown, L., and Pirkle, C.I. (1963) *Neisseria gonorrhoeae*. I. Virulence genetically linked to clonal variation. *J Bacteriol* **85**: 1274–1279.
- Kline, K.A., and Seifert, H.S. (2005a) Role of the Rep helicase gene in homologous recombination in *Neisseria gonorrhoeae*. *J Bacteriol* **187**: 2903–2907.
- Kline, K.A., and Seifert, H.S. (2005b) Mutation of the *priA* gene of *Neisseria gonorrhoeae* affects DNA transformation and DNA repair. *J Bacteriol* **187**: 5347–5355.
- Kline, K.A., Sechman, E.V., Skaar, E.P., and Seifert, H.S. (2003) Recombination, repair and replication in the pathogenic *Neisseriae*: the 3 R's of molecular genetics of two human-specific bacterial pathogens. *Mol Microbiol* **50**: 3–13.
- Knight, K.L. (1992) Restricted VH gene usage and generation of antibody diversity in rabbit. *Annu Rev Immunol* **10**: 593–616.
- Koomey, M., Gotschlich, E.C., Robbins, K., Bergstrom, S., and Swanson, J. (1987) Effects of *recA* mutations on pilus antigenic variation and phase transitions in *Neisseria gonorrhoeae*. *Genetics* **117**: 391–398.
- Koomey, M., Bergstrom, S., Blake, M., and Swanson, J. (1991) Pilin expression and processing in pilus mutants of *Neisseria gonorrhoeae*: critical role of Gly<sub>1</sub> in assembly. *Mol Microbiol* **5**: 279–287.
- Liang, F.T., Yan, J., Mbow, M.L., Sviat, S.L., Gilmore, R.D., Mamula, M., and Fikrig, E. (2004) *Borrelia burgdorferi* changes its surface antigenic expression in response to host immune responses. *Infect Immun* **72**: 5759–5767.
- Mehr, I.J., and Seifert, H.S. (1997) Random shuttle mutagenesis: gonococcal mutants deficient in pilin antigenic variation. *Mol Microbiol* **23**: 1121–1131.
- Mehr, I.J., and Seifert, H.S. (1998) Differential roles of homologous recombination pathways in *Neisseria gonorrhoeae* pilin antigenic variation, DNA transformation, and DNA repair. *Mol Microbiol* **30**: 697–710.
- Mehr, I.J., Long, C.D., Serkin, C.D., and Seifert, H.S. (2000) A homologue of the recombination-dependent growth gene, *rdgC*, is involved in gonococcal pilin antigenic variation. *Genetics* **154**: 523–532.
- Nash, H.A. (1996) Site-specific recombination: integration, excision, resolution, and inversion of defined DNA segments. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, Vol. 2. Neidhardt, F.C. (ed.). Washington, DC: American Society for Microbiology, pp. 2363–2376.
- Nassif, X., Beretti, J.L., Lowy, J., Stenberg, P., O'Gaora, P., Pfeifer, J., *et al.* (1994) Roles of pilin and PilC in adhesion of *Neisseria meningitidis* to human epithelial and endothelial cells. *Proc Natl Acad Sci USA* **91**: 3769–3773.
- Roberts, D.J., Craig, A.G., Berendt, A.R., Pinches, R., Nash, G., Marsh, K., and Newbold, C.I. (1992) Rapid switching to multiple antigenic and adhesive phenotypes in malaria. *Nature* **357**: 689–692.
- Rohrer, M.S., Lazio, M., and Seifert, H.S. (2005) A real-time, semi-quantitative RT-PCR assay demonstrates that the *pilE* sequence dictates the frequency and characteristics of pilin antigenic variation in *Neisseria gonorrhoeae*. *Nucleic Acids Res* **33**: 3363–3371.
- Roske, K., Blanchard, A., Chambaud, I., Citti, C., Helbig, J.H., Prevost, M.C., *et al.* (2001) Phase variation among major surface antigens of *Mycoplasma penetrans*. *Infect Immun* **69**: 7642–7651.
- Sechman, E.V., Rohrer, M.S., and Seifert, H.S. (2005) A genetic screen identifies genes and sites involved in pilin antigenic variation in *Neisseria gonorrhoeae*. *Mol Microbiol* **57**: 468–483.
- Segal, E., Billyard, E., So, M., Storzbach, S., and Meyer, T.F. (1985) Role of chromosomal rearrangement in *N. gonorrhoeae* pilus phase variation. *Cell* **40**: 293–300.
- Segal, E., Hagblom, P., Seifert, H.S., and So, M. (1986) Antigenic variation of gonococcal pilus involves assembly of separated silent gene segments. *Proc Natl Acad Sci USA* **83**: 2177–2181.

- Seifert, H.S. (1996) Questions about gonococcal pilus phase and antigenic variation. *Mol Microbiol* **21**: 433–440.
- Seifert, H.S. (1997) Insertionally inactivated and inducible *recA* alleles for use in *Neisseria*. *Gene* **188**: 215–220.
- Seifert, H.S., Wright, C.J., Jerse, A.E., Cohen, M.S., and Cannon, J.G. (1994) Multiple gonococcal pilin antigenic variants are produced during experimental human infections. *J Clin Invest* **93**: 2744–2749.
- Serkin, C.D., and Seifert, H.S. (1998) Frequency of pilin antigenic variation in *Neisseria gonorrhoeae*. *J Bacteriol* **180**: 1955–1958.
- Skaar, E.P., Lazio, M.P., and Seifert, H.S. (2002) Roles of the *recJ* and *recN* genes in homologous recombination and DNA repair pathways of *Neisseria gonorrhoeae*. *J Bacteriol* **184**: 919–927.
- Smith, J.D., Chitnis, C.E., Craig, A.G., Roberts, D.J., Hudson-Taylor, D.E., Peterson, D.S., *et al.* (1995) Switches in expression of *Plasmodium falciparum* var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* **82**: 101–110.
- Stohl, E.A., and Seifert, H.S. (2001) The *recX* gene potentiates homologous recombination in *Neisseria gonorrhoeae*. *Mol Microbiol* **40**: 1301–1310.
- Swanson, J., Bergstrom, S., Barrera, O., Robbins, K., and Corwin, D. (1985) Pilus– gonococcal variants. Evidence for multiple forms of piliation control. *J Exp Med* **162**: 729–744.
- Swanson, J., Bergstrom, S., Robbins, K., Barrera, O., Corwin, D., and Koomey, J.M. (1986) Gene conversion involving the pilin structural gene correlates with pilus+ in equilibrium with pilus– changes in *Neisseria gonorrhoeae*. *Cell* **47**: 267–276.
- Swanson, J., Robbins, K., Barrera, O., Corwin, D., Boslego, J., Ciak, J., *et al.* (1987) Gonococcal pilin variants in experimental gonorrhea. *J Exp Med* **165**: 1344–1357.
- Tu, Z.C., Wassenaar, T.M., Thompson, S.A., and Blaser, M.J. (2003) Structure and genotypic plasticity of the *Campylobacter fetus* sap locus. *Mol Microbiol* **48**: 685–698.
- Turner, C.M. (1997) The rate of antigenic variation in fly-transmitted and syringe-passaged infections of *Trypanosoma brucei*. *FEMS Microbiol Lett* **153**: 227–231.
- Wainwright, L.A., Pritchard, K.H., and Seifert, H.S. (1994) A conserved DNA sequence is required for efficient gonococcal pilin antigenic variation. *Mol Microbiol* **13**: 75–87.
- van der Woude, M.W., and Baumler, A.J. (2004) Phase and antigenic variation in bacteria. *Clin Microbiol Rev* **17**: 581–611.
- Zhang, Q.Y., DeRyckere, D., Lauer, P., and Koomey, M. (1992) Gene conversion in *Neisseria gonorrhoeae*: evidence for its role in pilus antigenic variation. *Proc Natl Acad Sci USA* **89**: 5366–5370.

### Supplementary material

The following supplementary material is available for this article online:

**Fig. S1.** Pilin variants arising from 1-81-S2 *pilE*.

This material is available as part of the online article from <http://www.blackwell-synergy.com>